Question 1

Youth trying cigarette is a concerning issue. It's worth investigating whether or not state level variance in the age they first try cigarettes outweigh that of school level variance.

In the model, i used weibull distribution since it is well suited for survival analysis. For school-level variation, i used loggamma(1,0.3) as prior. For state-level variation, i used loggamma(1,0.7) as prior. Figure 1 and 2 shows the prior and posterior distribution. Table 2 shows the mean of prior.

From Table 1, we can see state-level variance is higher than school-level variance. This supports the hypothesis. From Figure 3 we can see that the hazard is not flat. This implies children with same sex, ethnicity and rural/urban have different possibility of first try smoking in the upcoming month, regardless of their age.

In conclusion, by analyzing the data, we found supporting evidence for the first hypothesis, and evidence against the second one.

Table 1: School v.s. State level Variance

	mean	0.025quant	0.975quant
SD for school	0.172	0.148	0.200
SD for state	0.244	0.191	0.311

Table 2: Mean of Prior

	Mean
SD of School	0.241
SD of State	0.440

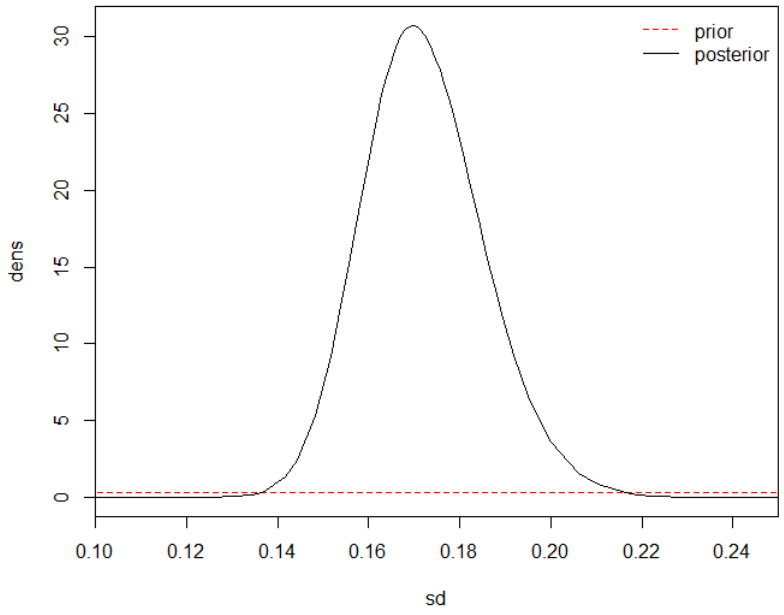

Figure 1: Prior and Posterior distribution of SD of School

Question 2

We are seeing increasing number of vehicles on the streets. Motor accidents become a major concern. Some car accidents do minor harm to the victim while some cause fatal injury. In this study we investigate in whether or not women are safer pedestrians than man.

We have done a case-control study. The data is divided into 3 strata. The light condition, the weather condition and the exact time the accident happened. A GLM and conditional logistic model is fit to the data.

From Figure 3, we can see that the negative estimate implies women on

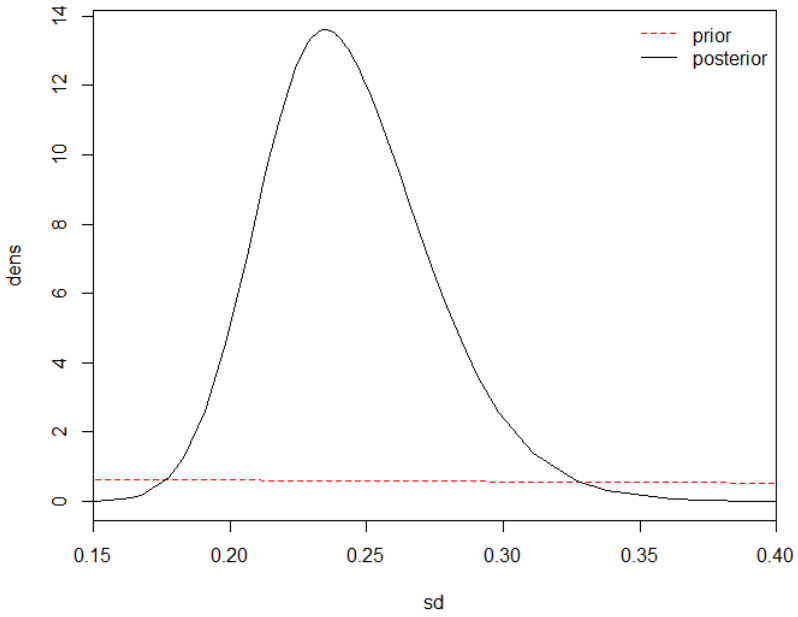

Figure 2: Prior and Posterior distribution of SD of State

average are less likely to have a fatal injury compared to man, therefore are safer pedestrians. From Figure 4, we can see that compared with men, women are much more safer until they reach upper twenties.

In conclusion, in the analysis we found supporting evidence for the hypothesis.

Table 3: Summary of the GLM Model

	Estimate	Std. Error	z value	P-value
(Intercept)	-3.251	0.024	-137.114	0.000
sexFemale	-0.299	0.012	-23.979	0.000
age0 - 5	0.112	0.034	3.255	0.001
age6 - 10	-0.437	0.032	-13.489	0.000

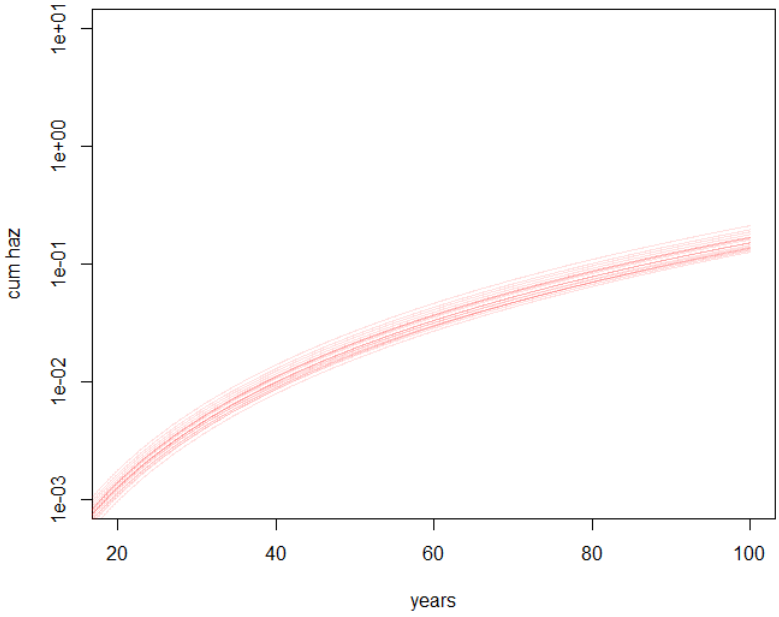

Figure 3: Cumulative hazard

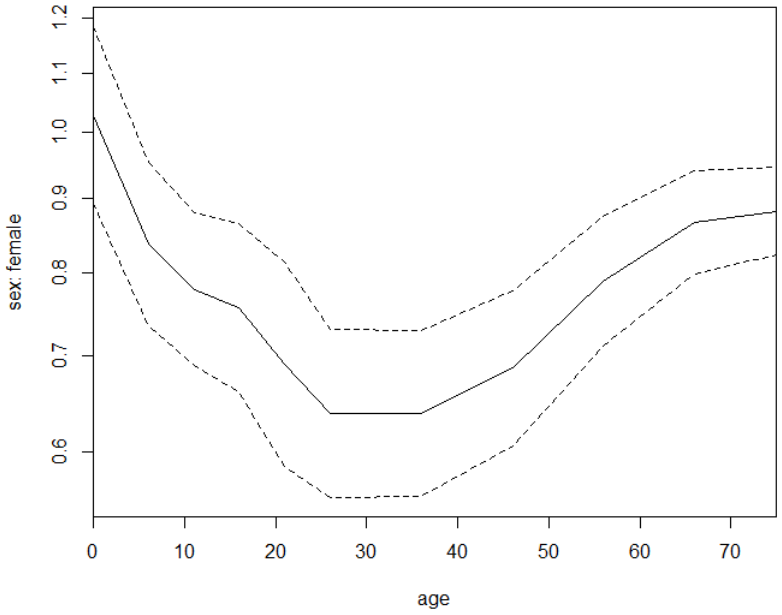

Figure 4: Women Relative to Man